NIH BACKGROUNDER

National Institutes of Health

Background on Genome-Wide Association Studies

December, 2006

Introduction

With the completion of the Human Genome Project in 2003 and the International HapMap Project in 2005, researchers have, for the first time, a set of research tools that make it possible to find the genetic contributions to common diseases such as heart disease, cancer and neurological disorders. The tools include computerized databases that contain the reference human genome sequence, a map of human genetic variation and a set of new lab techniques that can quickly, accurately and inexpensively analyze genetic samples for mutations that may contribute to the onset of a sickness. These new tools have led many of the 27 institutes and centers at the National Institutes of Health (NIH) to launch projects to find new genes or sets of genes that contribute to specific diseases. Once new gene targets are identified, researchers and companies can use the information to design effective new diagnostics, prevention strategies and treatments.

Genome-Wide Association Studies

This new approach is called genome-wide association because these studies rapidly scan the entire genome of a person and find changes that can be associated with a particular condition or disease. The design of a genome-wide association study, or GWAS, is conceptually easy to understand but challenging to conduct. Two groups of research subjects must be available: one group is made up of individuals suffering from the disease being studied and the other group is made up of people without the disease for comparison. Genetic samples are collected from both groups, either by drawing blood or rubbing a cotton swab along the side of the mouth to collect skin cells. The individuals' entire genome is then isolated from the samples and tested on an automated laboratory machine, of which several types are available from private companies. The machines read out any genetic differences, some of which may be associated with that disease.

The associated genetic differences themselves may not cause the disease. The associations, however, are often powerful pointers to that region of the human genome where the problem resides. Researchers then need to take additional steps to sort through the region to identify the exact genetic change and to understand its biological importance or function.

Researchers already have had a number of successes with this strategy, some producing surprising results. For example, in 2005, three independent studies reported finding an association between macular degeneration, a common form of blindness in old age, with a single mutation in a gene for complement factor H, which normally produces a protein involved in regulating inflammation. No one previously predicted that inflammation might contribute to this age-related blindness.

Genome-association studies produce massive sets of information for researchers to analyze. The National Center for Biotechnology Information, a part of NIH's National Library of Medicine, is

developing unique databases to make the results of GWAS studies widely and freely available for the research community. Two such studies on macular degeneration and Parkinson's disease already can be accessed at a newly created site called dbGaP (pronounced D, B, GAP) and that can be found at http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gap. NIH intends to make data from these studies available for individual researchers to zero in on the specific genetic changes that either cause or increase the susceptibility to these diseases.

The impact on medical care from GWAS will be substantial. Researchers predict that doctors will one day be able to cost-effectively use similar genome-wide scans of individual patients to predict for which disorders they are at increased risk. Doctors could then use that information to design prevention strategies to help the patient avoid becoming sick. In addition, knowing an individual's genotype should eventually allow doctors to customize medical treatment, selecting the most effective medication that will have the fewest undesirable effects. Experts already are calling this the era of personalized medicine in which the current one size-fits-all approach to treatment is replaced by customized therapies.

Examples of NIH Genome-Wide Association Studies

Numerous institutes at the National Institutes of Health have begun or will soon start studies to identify the genetic roots of common diseases..

In February, 2006, NIH and the Foundation for the National Institutes of Health (FNIH), announced a public-private partnership with Pfizer Global Research & Development, New London, Conn., to fund genome-wide association studies by the then newly launched Genetic Association Information Network (GAIN). In October, after competitive peer-review, GAIN announced the first round of studies would include bipolar disorders, depression and anxiety, kidney disease in type 1 diabetes, attention deficit hyperactivity disorder, schizophrenia and psoriasis. More information about GAIN can be found at http://www.fnih.org/GAIN/GAIN_home.shtml.

The Genes and Environment Initiative (GEI) is a new NIH funding program specifically designed to support genome-wide association studies. This $40 million initiative, proposed by Health and Human Services Secretary Michael O. Leavitt in the President's 2007 budget proposal, will create a pipeline for analyzing genetic variation in groups of patients with specific illnesses. The initiative adds a unique component that speeds up development of environmental exposure technology to produce new tools to study the interaction of genes and the environment. More information about GEI can be found at http://genesandenvironment.nih.gov/.

Individual institutes, too, have started GWAS studies. For example, the National Heart Lung and Blood Institute (NHLBI) launched the Framingham Genetic Research Study in collaboration with the Boston University School of Medicine in which the 9,000 study participants across three generations in the long-running Framingham Heart Study will undergo genome-wide association studies to identify the genes underlying cardiovascular and other chronic diseases. For more information on the study, see http://www.nhlbi.nih.gov/new/press/06-02-06.htm.

Some studies supported by NIH institutes already have been completed with data moving into the NCBI dbGaP website, including studies by the National Eye Institute on age-related eye diseases and the National Institute of Neurological Disorders and Stroke Parkinson's disease study.

Maximizing Public Benefit

In order to promote the greatest degree of scientific progress from GWAS funded with public support, an NIH-wide working group has been studying the important policy issues involved for much of 2006. Some of the questions to consider are: how to share data for research purposes with other qualified investigators, and, critically, how to protect the privacy of the people who participate in GWAS. The work of the NIH group led to the release of a proposed policy for GWAS over the summer on which NIH has been collecting public comments, including this town hall meeting. All comments received will be carefully considered by the NIH before issuing the final policy.